Facilitating Climate Research by Integrating NASA and the Earth System Grid

D. Crichton[1], Dean Williams[2], A. Braverman[1], Y. Chao[1], and C. Mattmann[1]
[1]Jet Propulsion Laboratory, Caltech; [2]Lawrence Livermore National Laboratory

Summary

Current climate researchers at NASA have an exceedingly difficult time finding and locating appropriate data sets and model outputs of use to their research. The existing process is extremely time-consuming, manual, labor-intensive work involving human-in-the-loop downloading and staging of (hundreds of) terabytes of distributed information. Our work will immediately increase data available to the climate modeling community's fingertips across NASA and other agencies, reducing the time to discover data by an order of magnitude over the current process. To do this, this project will connect NASA/Jet Propulsion Laboratory (JPL) into the DOE-funded Earth System Grid (ESG) and create an extensible software infrastructure to automatically retrieve and share data between NASA and the DOE. The ESG is a data grid featuring single sign-on access to massive (petabyte-scale) climate model runs accumulated at LLNL, NCAR and other leading climate research laboratories. ESG represents one of the most successful collaborations funded by the DOE Scientific Discovery Through Advanced Computation (SciDAC) program.

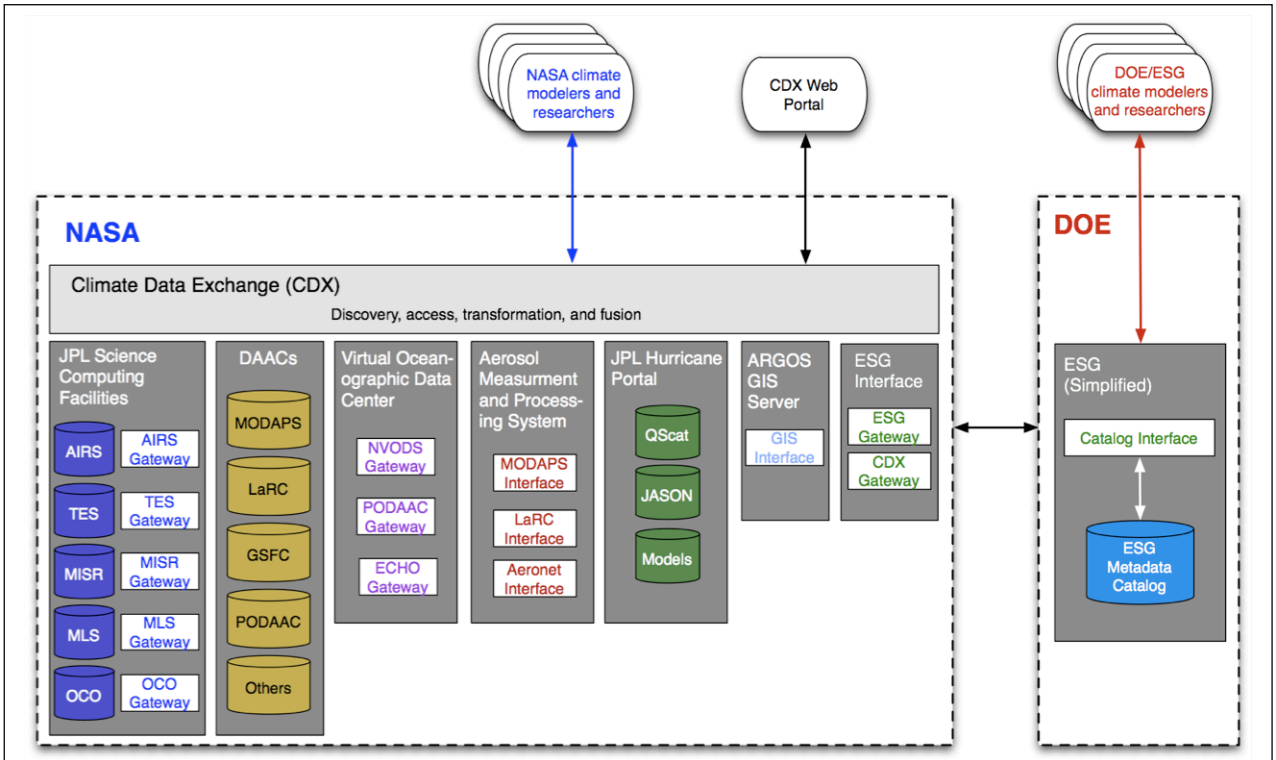

Figure 1: JPL's Climate Data Exchange for Connecting Models and Observational Data.

Global climate change is one of the great challenges and scientific questions facing our world today. Existing research and the promise of new investments are requiring information infrastructures to support new paradigms and approaches to conducting collaborative research at a national scale. Fused information products, correlative science, and the ability to perform distributed data analysis are all required in order to construct the predictive models that are needed to assess climate changes to the world in which we live. A major risk that currently exists is the disparity of data and technology across federal agencies in the area of climate research. Collaboration both within and across agencies is a necessity for advancing the state of knowledge.

Our primary objective is to enable distributed data analysis through sharing of observational data sets and climate model output data between NASA and DOE. Science research is dependent on obtaining access and sharing data that can be used to both discover and validate theories concerning greenhouse gases, water management, and changes to the ecosystem. NASA, as a pioneer in conducting Earth science research through sustained space-borne and airborne missions, has an enormous wealth of observational data that need to be used at a national level. Yet, a primary challenge is having the information infrastructure in place to share, transform and link together the necessary data products to support climate research.

The Department of Energy (DOE) has developed a sustained infrastructure through the Earth System Grid (ESG). It provides a number of information and science services that, if coupled with NASA data and capabilities, could create a multi-agency infrastructure for conducting climate research.

Both DOE and NASA have capabilities that can be directly leveraged to support and share model output and data between the two agencies. Within NASA, we will leverage the Climate Data Exchange (CDX) project. CDX utilizes JPL's Object Oriented Data Technology (OODT) middleware framework[1] for communication services, and common software interfaces (called *gateways*) to transparently access distributed data holdings and services. CDX has recently been connected to several science data systems and is being expanded to serve as an infrastructure for sharing data from JPL. It will include gateways to allow exchange of atmospheric Level 2 and Level 3 instrument data sets, data from the NASA Distributed Active Archive Centers (DAACs), as well as ESG/IPCC-provided climate model output (the ESG gateway) from Lawrence Livermore's[2] Program for Climate Modeling Diagnosis and Intercomparison (PCMDI).

From LLNL, this project will directly leverage and use both the modeling and software services built by the PCMDI. It will enable real-time access to PCMDI model outputs as well as support access to NASA observational data formatted using standard CF conventions.

The partnering arrangement between LLNL and JPL will be fully exploited to leverage the unique capabilities of each entity. LLNL will provide data and support for access to the holdings at PCMDI, particularly access to model output data. CDX will provide the software technology and support for building the interfaces to JPL observational instrument data systems. These will be brought together to begin intercomparison of PCMDI models to NASA observational data in order to support validation. To date, progress has already been made with NASA and DOE beginning to share data and connect their infrastructures.

NASA funding for this effort is supported through the Innovative Partnership Program (IPP).

For further information on this subject contact:
Name: Dan Crichton.
Organization: Jet Propulsion Laboratory
Email: Dan.Crichton@jpl.nasa.gov
Phone: (818) 354-9155

[1] Runner-up 2003 NASA Software of the Year
[2] Lawrence Livermore is a member of the Earth System Grid